Journal of Immunological Methods, 167 (1994) 279–287

JIM 06894

Two-stage selection of sequences from a random phage display library delineates both core residues and permitted structural range within an epitope

Robert M. Miceli *, Michael E. DeGraaf and H. David Fischer

Molecular Biology Research, The Upjohn Company, 301 Henrietta Street, Kalamazoo, MI 49007, USA

(Received 16 June 1993, revised received 1 September 1993, accepted 1 September 1993)

Libraries of random peptides can be screened to identify species which interact with antibodies or receptors. Similarly, maps of native molecular interactions can frequently be deduced by screening a limited set of peptide fragments derived from sequences within a native antigen or ligand. However, the existence of cross-reactive sequences that mimic original epitopes and the limited replaceability of amino acid residues suggest that the sequence space accessible by a receptor can be much broader. Definition of this space is of particular importance where structural information is required for peptidomimetic or drug design. We have used a two-stage selection scheme to expand the sequence space accessible by a phage display library and to define peptide epitopes of the anti-FLAG octapeptide monoclonal M2 antibody. Affinity selection of a primary library of 2×10^6 random decapeptides identified a non-contiguous core of three residues in the binding motif Tyr-Lys-Xaa-Xaa-Asp. A second stage library with 2×10^7 individual clones bearing the core motif but with the remaining flanking and internal residues re-randomized permitted access to a broader sequence space represented in a library equivalent to several orders of magnitude larger. Data here demonstrate that extended access to binding sequence space permitted by multi-stage screening of phage display libraries can reveal not only essential residues required for ligand binding, but also the ligand structural range permitted within the receptor binding pocket.

Key words: Epitope; Receptor; Peptide; Phage; Display library

Introduction

Vast numbers of ligands for antibodies or receptors can be screened by phage display as fusion proteins on the surface of filamentous bacteriophage (Scott and Smith, 1990; Cwirla et al., 1990). Successful use of these display libraries couples the expression of ligand-containing fusion proteins on the phage surface with amplification of ligand-encoding phage after affinity selection. Libraries of short random peptide sequences (Scott and Smith, 1990; Cwirla et al., 1990; Devlin et al., 1990), mutant and wild type hormones (Lowman et al., 1991), enzymes (Roberts et al., 1992; McCafferty et al., 1991) and antibody variable region genes (McCafferty et al., 1990; for review see Hoogenboom et al., 1992) have been successfully expressed on the surface of phage.

We chose the M2 monoclonal antibody/FLAG octapeptide epitope system (Hopp et al., 1988) to

* Corresponding author. Tel.: (616) 385-7455; Fax: (616) 385-5559.

SSDI 0022-1759(93)E0228-A

examine the selection and hierarchical dominance of peptides on phage for binding to a monoclonal antibody specific for the FLAG peptide immunogen. The FLAG peptide was originally engineered as a "tag" for immunoaffinity purification of genetically-engineered proteins. We identified two related peptide species after three rounds of affinity selection from an initial library of 2×10^6 random peptide phage clones. These selected peptides shared a common sequence motif with the original FLAG immunogen. While the sequence data obtained from the clones sampled following the initial selections provided preliminary positional information on critical core epitope residues, construction and screening of a secondary epitope library provided a more detailed map of the range of permissible residues within the epitope sequence. The theoretical sequence space accessible through the secondary library extended our primary library by seven orders of magnitude to over 10^{13} sequences, vastly greater than that contained in any single library reported to date (Cwirla et al., 1990; Scott and Smith, 1990; Devlin et al., 1990).

Materials and methods

Construction of the random decapeptide-pIII fusion phage display library

The fUSE-5 vector and associated *E. coli* host strains were generously provided by Dr. George Smith (University of Missouri, Columbia, MO) (Parmley and Smith, 1988; Scott and Smith, 1990). Construction of the random decapeptide display libraries has been described previously (DeGraaf et al., 1993). Briefly, a chemically synthesized, single-stranded degenerate insert was annealed to shorter oligonucleotides which generate *Sfi*I restriction site overhangs. A five-fold molar excess of the annealed DNA was mixed with 5 μg of *Sfi*I-cut fUSE-5 vector DNA and ligated overnight with T4 DNA ligase (Boehringer Mannheim). Following electroporation in *E. coli* MC1061 the library of 2×10^6 clones was amplified through approximately ten population doublings in fresh LB with 20 μg/ml tetracycline.

Construction of the secondary library of 2×10^7 individual clones with inserts encoding NH_2-Xaa-Xaa-Xaa-Tyr-Lys-Xaa-Xaa-Asp-Xaa-Xaa-COOH was performed in similar fashion to the primary library. Phage expressing the native FLAG peptide insert [$(Xaa)_3$-Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys-Gly] were also constructed and used for ELISAs as described below.

Affinity selection of phage libraries

Oxirane activated acrylic beads (Sigma) were washed extensively with PBS until the OD_{280} was approximately 0.00, followed by one wash with 1 M potassium phosphate buffer, pH 7.4. Coupling of anti-FLAG peptide monoclonal antibody (M2) (Kodak/IBI) was performed by incubating 100 μg antibody in 0.5 ml potassium phosphate buffer, pH 7.4 with 20 mg washed oxirane beads for 8 h at room temperature. Following coupling, the beads were reacted with 0.2 M β-mercaptoethanol for 4 h at 4°C to block excess oxirane groups and washed with TTB (50 mM Tris, pH 7.5, 150 mM NaCl, 0.5% Tween 20 (v/v), 1 mg/ml BSA) buffer.

The original random decapeptide phage library was concentrated to approximately 2×10^{10} pfu/ml in a 450 μl volume. A 1/20 vol. of this library was affinity selected on 20 mg of M2-coupled beads in TTB in a microfuge tube. Following 4 h incubation with rocking at room temperature the mixture was centrifuged at 14,000 rpm for 3 min and non-adherent phage in the supernatant were removed. Beads were washed five times with 1 ml TTB (15 min per wash) and bound phage eluted with 0.5 ml of 0.1 M glycine, pH 2.5 for 30 min at room temperature with gentle rocking. Following elution, the acid eluate was neutralized with 50 μl Tris-HCl, pH 9.5 and concentrated to 200–300 μl with a Centricon-30 microconcentrator (Amicon). Concentrated phage from the eluate were mixed with freshly-prepared "starved" *E. coli* K91-Kan cells and amplified overnight in LB containing 20 μg/ml tetracycline. Amplified phage were concentrated via sequential PEG precipitations as described (Parmley and Smith, 1988).

Biotinylation of epitope phage

Phage were equilibrated in 50 mM sodium bicarbonate, pH 8.5, diluted to a concentration of 2×10^{12} particles/ml and reacted with freshly

prepared sulfo-NHS-biotin (Pierce) at a final concentration of 500 μg/ml in the same buffer. The biotinylation reaction was allowed to proceed for 30–45 min at room temperature, quenched by the addition of solid glycine and unincorporated biotin removed by filtration through a Centricon-30 microconcentrator.

Phage ELISAs

Microtiter EIA II Plus plates (Linbro) were coated with 100 μl of FLAG MAb (6.0 μg/ml) in 0.1 M sodium bicarbonate, pH 8.5 at 4°C overnight. Plates were washed five times each with TBS/0.5% Tween 20 (v/v) and TBS before blocking with Blotto (5% nonfat dry milk (Carnation) in TBS, containing 0.02% NaN_3) for 45 min at 4°C. Plates were washed and biotinylated FLAG phage added (5×10^9 particles/well) in TBS/1% BSA. Soluble peptide competitors (native FLAG or variant FLAG peptides [133 uM]), were added during the phage binding step. Following binding for 45 min at 4°C, plates were washed and incubated for 30 min with avidin-HRP (Molecular Probes). Finally, plates were washed extensively, as above, and developed with *o*-phenylenediamine dihydrochloride (OPD) substrate (Sigma). Plates were read on an ELISA plate scanner (Titertek) at 450 nm after 5–10 min.

Results

Fig. 1 shows the sequences of decapeptide inserts in samples of phage selected from a primary library of 2×10^6 random clones through three rounds of affinity binding on immobilized M2 antibody. The data demonstrate a progressive selection of sequence species as a result of consecutive rounds of binding.

None of nine phage clones sampled after one round of binding selection was found in multiple copies, and seven of the nine were not seen again in samples from later rounds. One of the round

SELECTION ROUND	INSERT SEQUENCE	FREQUENCY	CLONE#
1	GLN-GLY-SER-ALA-**ASP-TYR-LYS**-TRP-SER-**ASP**	1/9	FR1.1
	PRO-HIS-**TYR-LYS**-ARG-GLY-ALA-ARG-LEU-**GLY**	1/9	FR1.2
	PHE-ILE-VAL-LEU-GLY-ASN-SER-**ASP**-PHE-TRP	1/9	FR1.3
	ILE-**LYS**-SER-VAL-SER-SER-ASN-**GLY**-PHE-SER	1/9	FR1.4
	LEU-LEU-ILE-**LYS**-ARG-ARG-GLY-SER-ASN-GLN	1/9	FR1.5
	PHE-ARG-SER-VAL-ALA-ARG-**LYS**-ALA-PHE-ARG	1/9	FR1.6
	LEU-GLY-ARG-**LYS**-LEU-PRO-GLU-HIS-LEU-SER	1/9	FR1.7
	ASP-ARG-SER-TRP-VAL-**TYR**- ? - ? - ?	1/9	FR1.8
	ILE-SER-LEU-SER-GLU-LEU-PRO-SER-GLU-ARG	1/9	FR1.9
2	LEU-**TYR-LYS**-GLU-GLY-**ASP**-VAL-PHE-ARG-SER	4/12	FR2.1
	GLN-GLY-SER-ALA-**ASP-TYR-LYS**-TRP-SER-**ASP**	6/12	FR2.2
	ALA-GLY-GLU-**TYR-LYS**-HIS-PHE-**ASP**-GLN-PHE	1/12	FR2.3
	TYR-LYS-ASP-GLY-**ASP**-HIS-TRP-PRO-ALA-VAL	1/12	FR2.4
3	LEU-**TYR-LYS**-GLU-GLY-**ASP**-VAL-PHE-ARG-SER	2/8	FR3.1
	GLN-GLY-SER-ALA-**ASP-TYR-LYS**-TRP-SER-**ASP**	6/8	FR3.2
	(ORIGINAL IMMUNOGEN)		
	-ASP-TYR-LYS-ASP-ASP-ASP-ASP-LYS-GLY-PRO-LYS-LYS-GLY-		
	1 2 3 4 5 6 7 8 9 10 11 12 13 (residue#)		

Fig. 1. Amino acid sequence of inserts deduced by DNA sequencing of phage clones following multiple rounds of affinity selection on M2 mAb-coated oxirane beads. Insert residues shown in bold show homology to the native FLAG peptide immunogen. Numbers in parentheses represent the number of repeated isolates/total number of clones sequenced at that round.

one clones, FR1.2, contained a TYR-LYS dipeptide shared with the original immunogen, but did not become enriched by subsequent rounds of affinity selection. By contrast, the remaining first round clone (FR1.1) was found duplicated (by clone FR2.2) as half of the 12 clones sampled after round two and 6 (clone FR3.2) of the 8 clones picked after three selection cycles. The Asp-Tyr-Lys-Xaa-Xaa-Asp sequence motif of the decapeptide insert in these highly-enriched clones also had the most similarity of any of the sampled sequences to the original FLAG immunogen. The high frequency of representation with repeated selection, combined with the early appearance of this sequence after just one round of selection suggests its strong, specific enrichment by affinity.

A second clone (FR2.1) bearing the Tyr-Lys-Xaa-Xaa-Asp motif in FLAG was similarly multiply represented in one third of the clones at round two and was identical to the less frequent (FR3.1) of the two unique phage sequences sampled after round three. Each of the other two clones in the sample after two selection rounds also shared the Tyr-Lys-Xaa-Xaa-Asp motif of the immunogen in different overall sequences, but were each present only in single copies and only at that round. One of these clones (FR2.4) also contained an additional Asp residue in common with position 4 of the FLAG immunogen. This added sequence homology did not, however, appear to confer a selective advantage over other identified species in subsequent rounds of binding. This clone was not present in multiple copies and its selection was not sustained through further enrichment in subsequent rounds. This contrasts with the early dominance of clones bearing an Asp residue immediately preceding the Tyr-Lys dipeptide at immunogen position 1, which suggested a possible binding advantage to ligands with this residue.

The importance of the shared Asp-Tyr-Lys-Xaa-Xaa-Asp motif for M2 antibody binding was confirmed by competitive ELISAs with phage deliberately engineered to bear the native FLAG immunogen sequence. The data in Fig. 2 show that binding of the FLAG peptide phage by M2 antibody could be inhibited by a synthetic peptide with the same immunogen sequence. Similarly, a synthetic peptide with the Asp-Tyr-Lys-Xaa-Xaa-Asp core sequence of the selected peptide phage (but containing neither of the two internal Asp residues present at positions 4 and 5 in the original immunogen) also blocked binding of the native FLAG phage. The high concentration of soluble peptide required (133 μM) to inhibit binding by multivalent phage is consistent with reports by others (Barrett et al., 1992). By contrast, a synthetic peptide with no homology to the FLAG immunogen, and a peptide sequence from a phage containing only the Tyr-Lys dipeptide of the identified core motif failed to inhibit binding of the native peptide phage.

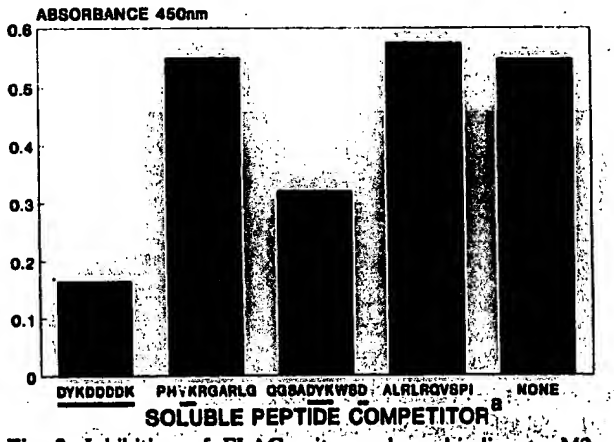

Fig. 2. Inhibition of FLAG-epitope phage binding to M2 antibody by soluble synthetic peptide competitors. Underlined letters in peptide sequence represent shared residues with the native FLAG immunogen. [a] final peptide concentration 133 μM.

More detailed resolution of the critical epitope residues flanking and within the putative core motif was obtained with a secondary library bearing $(Xaa)_3$-Tyr-Lys-Xaa-Xaa-Asp-$(Xaa)_2$ inserts randomized at non-core amino acids. An Asp corresponding to position #1 of the FLAG immunogen in the core motif was not maintained here as a constant residue. It's presence in the sequences selected from the initial library was confined to only a single sequence in sibling clones and could dominate for reasons other than optimal binding. Selection by affinity binding to immobilized M2 antibody was done identically to that of the primary library.

Data in Fig. 3 show that no repeats were found in the sequences of eight clones sampled after

ORIGINAL IMMUNOGEN:				-ASP-TYR-LYS-ASP-ASP-ASP-ASP-LYS-					
RESIDUE POSITION:	(-2)	(-1)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
SELECTION ROUND	INSERT SEQUENCE								
1	-lys-ser-ASP-TYR-LYS-glu-gln-ASP-arg-thr-								
	-lys-ser-ASP-TYR-LYS-glu-gln-ASP-arg-thr-								
	-glu-asp-ASP-TYR-LYS-asn-glu-ASP-ile-LYS-								
	-val-asn-ASP-TYR-LYS-ala-pro-ASP-thr-arg-								
	-leu-ile-ASP-TYR-LYS-cys-arg-ASP-ser-thr-								
	-leu-ser-ASP-TYR-LYS-gly-ala-ASP-arg-ile-								
	-gly-leu-ASP-TYR-LYS-ASP-gly-ASP-pro-cys-								
	-val-ser-ser-TYR-LYS-ASP-ASP-ASP-gln-ile-								
	-phe-glu-met-TYR-LYS-ASP-leu-ASP-thr-arg-								
	-ile-ile-gly-TYR-LYS-ASP-ile-ASP-gly-ser-								
	-arg-ala-phe-TYR-LYS-ASP-gly-ASP-pro-ser-								
	-arg-glu-glu-TYR-LYS-ASP-glu-ASP-gly-ile-								
	-phe-ala-glu-TYR-LYS-met-gly-ASP-leu-ile-								
	-met-val-glu-TYR-LYS-trp-ala-ASP-phe-met-								
	-leu-arg-glu-TYR-LYS-met-leu-ASP-ala-tyr-								
	-ile-lys-asn-TYR-LYS-phe-met-ASP-pro-thr-								
2	-pro-val-ASP-TYR-LYS-leu-ASP-ASP-ASP-ile								
	-ala-val-ASP-TYR-LYS-val-ASP-ASP-leu-leu-								
	-ile-ile-ASP-TYR-LYS-ser-ser-ASP-phe-ser-								
	-lys-gly-ASP-TYR-LYS-ala-phe-ASP-asn-leu-								
	-arg-ser-ASP-TYR-LYS-ala-gly-ASP-thr-ser-								
	-arg-ser-ASP-TYR-LYS-phe-arg-ASP-leu-ser-								
	-arg-asp-ASP-TYR-LYS-lys-gly-ASP-gln-tyr-								
	-leu-met-ASP-TYR-LYS-phe-ser-ASP-lys-tyr-								
	-gly-ser-ASP-TYR-LYS-leu-arg-ASP-pro-gly-								
	-val-lys-ASP-TYR-LYS-tyr-gly-ASP-ile-lys-								
	-arg-ile-ASP-TYR-LYS-arg-val-ASP-ala-phe-								
	-asn-gly-ASP-TYR-LYS-arg-ser-ASP-thr-met-								
	-arg-thr-ASP-TYR-LYS-phe-val-ASP-tyr-thr-								
	-val-gln-ASP-TYR-LYS-gly-ala-ASP-arg-gly-								
	-trp-asn-asn-TYR-LYS-ASP-phe-ASP-glu-leu-								
	-glu-ile-ser-TYR-LYS-phe-gln-ASP-pro-thr-								
3	-his-asn-ASP-TYR-LYS-leu-val-ASP-gln-arg-								
	-asp-lys-ASP-TYR-LYS-asn-val-ASP-gln-leu-								
	-ile-ser-ASP-TYR-LYS-arg-ala-ASP-lys-ser-								
	-asn-val-ASP-TYR-LYS-phe-tyr-ASP-lys-ala-								
	-leu-arg-ASP-TYR-LYS-ile-ala-ASP-tyr-tyr-								
	-ser-tyr-ASP-TYR-LYS-gln-ser-ASP-arg-val-								
	-asp-pro-ASP-TYR-LYS-ala-ala-ASP-val-gly-								
	-his-val-ASP-TYR-LYS-trp-pro-ASP-leu-his-								
	-cys-val-ASP-TYR-LYS-gln-val-ASP-lys-ser-								
	-arg-lys-ASP-TYR-LYS-leu-his-ASP-phe-leu-								
	-leu-arg-ASP-TYR-LYS-glu-leu-ASP-val-phe-								
	-ser-gly-ASP-TYR-LYS-ASP-lys-ASP-thr-arg-								
	-gly-arg-ASP-TYR-LYS-tyr-ASP-ASP-thr-arg-								
	-asp-val-asn-TYR-LYS-ASP-leu-ASP-ile-phe-								
	-phe-tyr-glu-TYR-LYS-leu-ala-ASP-ala-leu-								
	-phe-ile-glu-TYR-LYS-leu-gly-ASP-ile-pro-								
4	-gly-lys-ASP-TYR-LYS-ile-ala-ASP-leu-asn-								
	-gly-lys-ASP-TYR-LYS-ile-ala-ASP-leu-asn-								
	-gly-lys-ASP-TYR-LYS-ile-ala-ASP-leu-asn-								
	-arg-ser-ASP-TYR-LYS-arg-ser-ASP-arg-leu-								
	-arg-ser-ASP-TYR-LYS-arg-ser-ASP-arg-leu-								
	-ile-val-ASP-TYR-LYS-met-arg-ASP-ser-arg-								
	-ile-val-ASP-TYR-LYS-met-arg-ASP-ser-arg-								
	-leu-val-ASP-TYR-LYS-ala-ser-ASP-ala-arg-								
	-arg-gln-ASP-TYR-LYS-leu-val-ASP-arg-asp-								
	-tyr-thr-ASP-TYR-LYS-lys-his-ASP-lys-gly-								
	-leu-val-ASP-TYR-LYS-thr-ser-ASP-ser-phe-								
	-arg-asp-ASP-TYR-LYS-ala-phe-ASP-pro-arg-								
	-leu-ser-ASP-TYR-LYS-his-glu-ASP-arg-trp-								
	-trp-ile-ser-TYR-LYS-ASP-gly-ASP-thr-arg-								
	-val-ala-ser-TYR-LYS-leu-leu-ASP-pro-ser-								

Fig. 3. Amino acid sequence of inserts deduced by sequencing of secondary library clones selected through multiple rounds of binding to immobilized M2 antibody. Amino acid positions are numbered relative to corresponding residues in the original FLAG peptide immunogen.

one round of selection. Several of the selected clones, however, did bear additional homology to the original immunogen with an Asp residue at either the first or fourth position of the FLAG sequence. One of these also shared an Asp residue with the native immunogen at position five.

Further rounds of affinity selection with the secondary library revealed several patterns. First, the selected sequences were highly heterogeneous at positions 4 and 5 within the core motif of the immunogen. This result is shown in detail in Fig. 4. 19 and 16 of the possible 20 amino acids were represented at the respective positions among the total of 63 clones sampled across four consecutive rounds of selection. Sequences sharing an Asp residue at either of these two positions with the original immunogen did not dominate the selected clones (Figs. 3 and 4). Similarly, the residues at position 7 were similarly heterogeneous and represented by seventeen of the possible twenty amino acid residues (Figs. 3 and 4).

By contrast, Fig. 3 does show preferences upon successive selection toward acidic residues Glu (in six of 63 clones) or Asp (47 of 63) at immunogen position 1. As shown in Fig. 5, the frequency of Asp residues occurring at position 1 in the secondary library was significantly increased through several rounds of repetitive selection on the M2 antibody. By the fourth round of selection, 13 out of 15 sampled sequences included the position 1 Asp. No dominance for the other Asp residues of the native FLAG were observed at any of the other rerandomized positions.

			LYS				
			TYR				
			VAL			TYR	
			TYR	LYS		LYS	
			GLN	HIS		GLU	
			ILE	TYR		ASN	
			LEU	VAL		LEU	
			SER	PRO		ILE	
			ALA	GLU		THR	
			ASN	ARG		GLY	
PHE			GLY	MET		PRO	
MET			CYS	ILE		ALA	
ASN			PHE	LEU		ARG	
GLY			MET	PHE		PHE	
SER			TRP	SER		HIS	
HIS			ARG	ALA		SER	
GLU			HIS	GLN		GLN	
LEU			GLY	GLY		VAL	
-ASP-	-TYR-	-LYS-	-ASP-	-ASP-	-ASP-	-ASP-	
	*	*			*		

Fig. 4. Positional heterogeneity within the FLAG epitope reflected by secondary library clones. Data compiled from analysis of the 63 clones shown in Fig. 3 selected through four rounds of successive selection on M2 antibody. Asterisks denote fixed residues within the secondary library insert.

While the data clearly suggest a selective process at work during the multiple rounds of binding, no one individual clone from the secondary library dominated even after four rounds, as might be anticipated from the much larger representation of expected binding peptides synthesized around the core binding motif. Hawkins et al. (1992) derived a model for affinity selection of phage display clones that shows enrichment of phage with higher affinity depends not only on the difference in affinity between phage species, but also the concentration of receptor relative to the binding constant. They showed relative enrichment by affinity in solution phase depends on the receptor concentration being less than the K_d for receptor-ligand binding. This can be achieved in one way with low target receptor concentrations in solution or fewer immobilized receptors in solid phase. We have chosen instead to mimic these conditions (Chaiken et al., 1989) by retaining a high absolute number of immobilized receptors but increasing the apparent (functional) K_d of the selective interaction with non-chaotropic, non-specific inhibitors in the binding solution (Thompson et al., 1992). ELISA data (not shown) demonstrated that these conditions significantly decreased binding of several phage isolates while

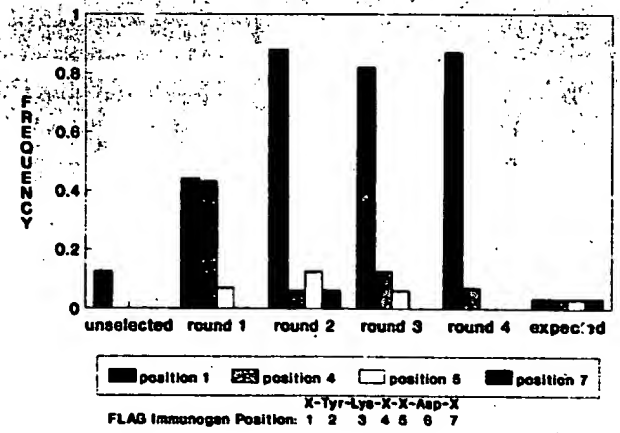

Fig. 5. Frequency of Asp residues in selected secondary library clones. Residue poistions are numbered relative to residue positions in the original FLAG peptide immunogen. Unselected frequency represents the occurrence of Asp at each position in eight clones sequenced prior to selection. "Expected" values represent stochastic frequency of occurrence for Asp dictated by genetic code (2/61).

INSERT SEQUENCE	FREQUENCY
-tyr-ile-ASP-TYR-LYS-*val-phe*-ASP-met-phe-	6/15
-ala-lys-ASP-TYR-LYS-*leu-phe*-ASP-pro-leu-	4/15
-val-lys-ASP-TYR-LYS-*leu*-val-ASP-his-leu-	1/15
-ser-asp-ASP-TYR-LYS-*val*-gln-ASP-tyr-phe-	1/15
-ala-gly-ASP-TYR-LYS-*leu*-ile-ASP-gly-pro-	1/15
-leu-tyr-ASP-TYR-LYS-asn-gln-ASP-trp-gly-	1/15
-trp-arg-thr-TYR-LYS-glu-tyr-ASP-val-gly-	1/15
-ASP-TYR-LYS-ASP-ASP-ASP-ASP-LYS-	FLAG IMMUNOGEN
(-2) (-1) (1) (2) (3) (4) (5) (6) (7) (8)	RESIDUE #

Fig. 6. Peptide sequences selected from inserts in the secondary library following one round of affinity selection under stringent binding conditions. The sequences are aligned relative to identical residues in the original FLAG peptide immunogen. Stringent selection conditions were similar to those described in Materials and Methods except that ethylene glycol (40% v/v in TTB) and salt (to a final concentration of 320 mM) were added during the binding step and to the washes immediately following. Frequency represents number of repeated isolates/total number of clones sequenced.

others, including the native FLAG sequence, were minimally affected.

As shown above, dominant representation of Asp residues corresponding to the first position of the FLAG immunogen required two to three rounds of selection by standard conditions. By contrast, Fig. 6 shows that 14 of 15 sampled phage clones bore an Asp residue at position 1 following just one round of selection in buffers with added ethylene glycol and increased salt. This almost uniform representation is contributed by six distinct phage sequences – each displaying an Asp residue at that first position. Two of the clones are represented as apparently sibling clones in multiple copies. In addition, the clones selected under the modified conditions displayed a bias toward hydrophobic amino acids in the residues at positions 4 and 5 of the immunogen core sequence. The predominance of Val (7/15) and Leu (6/15) at position 4 and Phe (10/15) at position 5 suggest an affinity-driven selection for hydrophobic residues at these positions.

Discussion

Our use of a two-stage selection using modestly sized phage display libraries (ca. 2×10^6–2

$\times 10^7$ independent clones) has demonstrated (1) the identification of key residues involved in the M2 epitope core and (2) a permissible range of flanking and intervening epitope residues.

Successive cycles of affinity-based selection of a random decapeptide phage library yielded sequence information sufficient to predict a core epitope for the M2 monoclonal antibody. The relative importance of key core residues could be predicted through the dominance of individual sequences after multiple rounds of affinity binding and the pattern of sequences which emerged through consecutive rounds of selection.

The data from the primary library suggested a core epitope with three or four key residues spanning six amino acid positions. A complete library of peptides with all twenty amino acids represented at each of six residue positions comprises 6.4×10^7 different species. One strategy for accessing extended sequence space with combinatorial libraries has focused on making larger starting libraries. Here we describe a different approach exploiting principles used by the immune system. First, a diverse but limited repertoire of peptide binding species (DeGraaf et al., 1993) analogous to the available antibody repertoire represented on primary $sIgM^+$ B cells provides a starting point for defining and ultimately strengthening the molecular interactions of ligand and receptor. Second, species with preferred binding are enriched by affinity selection in a process comparable to the clonal selection and expansion of primary B cells. Third, basic sequence and structural themes of selected receptor-bearing clones (in the case of B cells) or candidate ligands (in the case of display peptides) are then modified through mutation. Through cumulative selection building on a common, but selected theme, repertoires of modest size can be expanded to provide adequate information for mapping of flanking epitope residues and analysis of residue structural constraints. Thus, the need to construct larger, logistically unmanageable libraries becomes less necessary.

In this work, two-stage selection provided two types of structural information important for molecular modeling and epitope mapping from combinatorial libraries. First, sequences selected from a small, primary library delineated key, core residues required for M2 receptor binding. Second, the diversity provided by a secondary library provided access not only to the key species, but also to a range of permitted species. This approach is conceptually similar to that described earlier (Geysen et al., 1986), where an identified di- or tripeptide core of an epitope is the basis for chemical synthesis of broader synthetic peptide sets screened to extend knowledge of the epitope and surrounding residues. We have incorporated this approach to epitope mapping into the phage display system by construction of a secondary library. While we have demonstrated this method of ligand assessment on a continuous epitope, it should also prove to be a valuable strategy in defining spatial constraints in peptide mimics of discontinuous ligands as well. Modeling based on the range of amino acid types permissible at each position within a preferred ligand could help determine the size and contour of a receptor binding pocket for analog engineering and drug discovery.

The importance of the predicted Asp-Tyr-Lys-Xaa-Xaa-Asp core motif in M2 binding was confirmed by independent solid-phase binding assays. Further, the broad diversity of structure observed at the two intervening positions within the core motif argues strongly for the lack of steric and spatial restrictions at those sites. Recent evidence by Gong et al. (1993) show a high degree of correlation between residues selected from a random epitope library and those shown to be critical by crystallographic analysis. Similarly, we have used NMR spectroscopy of [^{15}N]-D_1YKLGD_6D_7L (99 atom % ^{15}N at the α position of each amino acid except for lysine) to investigate the interactions of this peptide with M2 Fab fragments. The largest chemical shift differences between the free and bound peptide were observed for residues D_1 and D_6, indicating that these residues are those most likely to interact with the Fab fragment (Stockman et al., manuscript in preparation). These independent observations support our predictions as to the importance of Asp residues at positions 1 and 6 of the immunogen in the M2 epitope. Both Asp residues were selected in either primary or secondary libraries (the position 6 Asp residue was fixed in the secondary library as a core residue).

The immune system exploits relative binding affinity to select and amplify preferred structural variations of an originally selected clone. We believe we can functionally mimic the latter with peptide phage through binding kinetics. In our experiments the higher stringency conditions during selection appeared to enrich for selected species from a larger background of specific, but functionally less efficacious binders. Our choice and use of buffer conditions here was aided by our prior knowledge of the original immunogen and our ability to directly test buffer effects on its antibody binding. For an unknown ligand where the binding mechanism is not known, buffer conditions might first be tested on candidate species selected from the primary library. Otherwise, arbitrary modulation of binding conditions such as hydrophobicity or ionic strength might functionally bias selection to specific peptide types rather than necessarily the best overall binders. This approach could, in fact, be exploited practically as an additional selection condition for ligand species in specific structural classes or with desired binding properties.

Since there is little chance that a phage display library can be constructed that fully represents all peptide ligands longer than about six or seven residues, multistage selections like that described here can provide one means to effectively explore a much fuller sequence space. Such a search strategy depends first on broad diversity in the sequence space coverage of the primary library (DeGraaf et al, 1993), and on selection protocols with sufficient resolution to discriminate a small number of strong, optimal binders from a background of statistically more numerous, but less effective ligands.

Finally, when building upon or modifying sequences identified in the first-stage screening of a primary library, it is important to recognize that the selected core sequences represent individual, local optima in the sequence space that was available and effectively searched (Kauffman, 1993). A global optimum might be missing or under-represented in any discontinuous library. This fact can effect the method chosen to modify initially-selected core species to explore a larger sequence space (Kauffman, 1993). Rather than maintain complete constancy of core residues as in the example reported here, alternative mutation strategies, including codon-based mutagenesis (Glasser et al., 1992) or an error-prone DNA polymerase chain reaction (Pannekoek et al., 1993) might be employed to increase diversity in regions around such local optimum while also maintaining a cumulative selection upon the initial sequence motif. Alternatively, a larger panel of candidate core sequences, or a less complete core sequence might be chosen for mutagenesis instead of the single motif selected here.

Acknowledgements

The authors acknowledge the excellent technical help of Sarah J. Neering and thank Kathy Kronke for help in the preparation of this manuscript.

References

Barrett, R.W., Cwirla, A.E., Ackerman, M.S., Olson, A.M., Peters, E.A. and Dower, W.J. (1992) Selective enrichment and characterization of high affinity ligands from collections of random peptides on filamentous phage. Anal. Biochem. 204, 357.

Chaiken, I., Fassina, G. and Shai, Y. (1989) Analytical affinity chromatography, antisense peptides and the synthetic design of recognition molecules. In: T.W. Hutchens (Ed.), Protein Recognition of Immobilized Ligands. Alan R. Liss, New York, p. 3.

Cwirla, S.E., Peters, E.A., Barrett, R.W. and Dower, W.J. (1990) Peptides on phage: A vast library of peptides for identifying ligands. Proc. Natl. Acad. Sci. USA 87, 6378.

Devlin, J.J., Panaganiban, L.C. and Devlin, P.E. (1990) Random peptide libraries: A source of specific protein binding molecules. Science 249, 404.

DeGraaf, M.E., Miceli, R.M., Mott, J.E. and Fischer, H.D. (1993) Biochemical diversity in a phage display library of random decapeptides. Gene 128, 13.

Geysen, H.M., Rodda, S.J. and Mason, T.J. (1986) A priori delineation of a peptide which mimics a discontinuous antigenic determinant. Mol. Immunol. 23, 709.

Glasser, S.M., Yelton, D.E. and Huse, W.D. (1992) Antibody engineering by codon-based mutagenesis in a filamentous phage vector system. J. Immunol. 149, 3903.

Gong, X., Schopp, C. and Scott, J.K. (1993) The performance of epitope libraries: Crystallographic evidence for a direct correspondence between the residues selected and the critical residues for binding. Protein Eng. 6, 106.

Hawkins, R.E., Russell, S.J. and Winter, G. (1992). Selection of phage antibodies by binding affinity: Mimicking affinity maturation. J. Mol. Biol. 226, 889.

Hoogenboom, H.R., Marks, J.D., Griffiths, A.D. and Winter, G. (1992) Building antibodies from their genes. Immunol. Rev. 130, 41.

Hopp, T.P., Prichett, K.S., Price, V.L., Libby, R.T., March, C.J., Cerretti, D.P., Urdal, D.L. and Conlon, P.J. (1988) A short polypeptide marker sequence useful for recombinant protein identification and purification. Bio/Technology 6, 1204.

Kauffman, S.A. (1993) The structure of adaptive landscapes underlying protein evolution. In: S.A.Kauffman, The Origins of Order: Sel-organization and Selection in Evolution. Oxford University Press, Oxford, p. 121.

Lowman, H.B., Bass S.H., Simpson, N. and Wells, J.A. (1991) Selecting high-affinity binding proteins by monovalent phage display. Biochemistry 30, 10832.

McCafferty, J., Griffiths, A.D., Winter, G. and Chiswell, D.J. (1990) Phage antibodies: filamentous phage display of antibody variable domains. Nature 348, 552.

McCafferty, J., Johnson, R.H. and Chiswell, D.J. (1991) Phage-enzymes: Expression and affinity chromatography of functional alkaline phosphatase on the surface of a bacteriophage. Protein Eng. 4, 955.

Pannekoek, H., van Meijer, M., Schleef, R.R., Loskutoff, D.J. and Barbas, C.F. (1993) Functional display of human plasminogen-activator inhibitor 1 (PAI-1) on phages: novel perspectives for structure-fucntion analysis by error-prone DNA synthesis. Gene 128, 135.

Parmley, S.F. and Smith, G.P. (1988) Antibody-selectable filamentous fd phage vectors: Affinity purification of target genes. Gene 73, 305.

Roberts, B., Markland, W., Ley, A.C., Kent, R.B., White, D.W., Gutermen, S.K. and Ladner, R.C. (1992) Directed evolution of a protein: Selection of potent neutrophil elastase inhibitors displayed on M13 fusion phage. Proc. Natl. Acad. Sci. USA 89, 2429.

Scott, J.K. and Smith, G.P. (1990) Searching for peptide ligands with an epitope library. Science 249, 386.

Thompson, N.E., Hager, D.A. and Burgess, R.R. (1992) Isolation and characterization of a polyol-responsive monoclonal antibody useful for gentle purification of *E. coli* RNA polymerase. Biochemistry 31, 7003.